

ABSTRAK

RAMLI SALASA: Komparasi Metode Penyetaraan Vertikal pada Skor Mata Pelajaran Matematika Siswa SMP Negeri 1 Tidore Kepulauan. **Tesis. Program Pascasarjana, Universitas Negeri Yogyakarta, 2014.**

Penelitian ini bertujuan untuk: (1) mengetahui karakteristik butir soal berdasarkan hasil analisis dengan menggunakan teori tes klasik, (2) mengetahui karakteristik butir soal berdasarkan hasil analisis dengan menggunakan teori respon butir dengan model satu parameter (*Model Marginal Maximum Likelihood*), (3) membandingkan hasil analisis butir soal dengan menggunakan teori tes klasik dan teori respon butir, (4) mengetahui kemampuan peserta ujian, dan (5) mengetahui persamaan penyetaraan perangkat tes prestasi belajar Matematika antara kelas VII, VIII, dan IX SMP Negeri 1 Tidore Kepulauan.

Penelitian ini dilakukan dengan metode kuantitatif. Populasi penelitian ini adalah siswa SMP Negeri 1 Tidore Kepulauan, dengan jumlah total sebanyak 606 siswa. Dalam penyetaraan butir-butir yang digunakan dimasukkan ke dalam bank soal adalah butir-butir yang telah tersaring melalui analisis butir, baik klasik maupun modern. Estimasi parameter dilakukan berdasarkan model logistik 2 parameter dari dalam analisisnya menggunakan program komputer BILOG 3.00. Data yang berupa skor diolah dan disetarakan dengan menggunakan pendekatan teori tes klasik (*Microcat iteman*) dan teori tes modern.

Hasil penelitian menunjukkan bahwa soal-soal Matematika pada ujian Matematika kelas VII, VIII, dan IX sudah cukup baik kualitasnya. Berdasarkan tes klasik perangkat tes kelas VII sejumlah 29 butir, perangkat tes kelas VIII sejumlah 27 butir, dan perangkat tes kelas IX terdiri 30 butir. Sedangkan berdasarkan tes modern perangkat tes VII sejumlah 28 butir, perangkat tes kelas VIII sejumlah 27 butir, dan perangkat tes kelas IX terdiri 24 butir. Penyetaraan perangkat tes butir soal dari kelas VII ke kelas VIII adalah $Q^* = 1,084X - 0,221$, dari kelas VIII ke kelas IX adalah $R^* = 0,922X + 2,965$ dan dari kelas VII ke IX adalah $R'^* = 0,998 X + 2,775$. Kemampuan peserta ujian Matematika kelas VII, VIII, dan IX menunjukkan bahwa rata-rata kemampuan siswa berkisar antara 0,265 sampai dengan 0,58. Artinya kemampuan peserta ujian kelas masih rendah. Penyetaraan perangkat tes kelas VII ke kelas VIII mengalami penurunan, menjadi lebih mudah bila dikerjakan oleh siswa kelas VIII. Begitu pula penyetaraan perangkat tes kelas VIII ke kelas IX mengalami penurunan, menjadi lebih mudah bila dikerjakan oleh siswa kelas IX. Namun Penyetaraan perangkat tes kelas VII ke kelas IX mengalami peningkatan, menjadi lebih sulit bila dikerjakan oleh siswa kelas IX. Hal ini bisa disebabkan karena kelas IX kurang memahami kembali pelajaran kelas VII. Berdasarkan uraian analisis tes menggunakan dua pendekatan yaitu tes klasik dan tes modern (*Model Marginal Maximum Likelihood*) dapat ditunjukkan bahwa dalam kasus mata pelajaran Matematika pada ketiga kelas yang diamati terjadi perbedaan dalam persentase validitas butir yang baik.

Kata Kunci: *penyetaraan vertikal, tes prestasi belajar, matematika, bank soal*

ABSTRACT

RAMLI SALASA: *A Comparison of Vertical Equating Methods for Mathematics Scores of the Students of SMP Negeri 1 Tidore Kepulauan.* Thesis. Yogyakarta: Graduate School, Yogyakarta State University, 2014.

This study aims to: (1) investigate characteristics of test items based on the results of the analysis using classical test theory, (2) investigate characteristics of test items based on the results of the analysis using item response theory with the one parameter model (the Marginal Maximum Likelihood model), (3) compare the results of the item test analyses using classical test theory and item response theory, (4) investigate testees' ability, and (5) investigate equations for equating test sets for mathematics learning achievement for Grades VII, VIII, and IX of SMP Negeri 1 Tidore Kepulauan.

The study was conducted using the quantitative method. The research population consisted of the students of SMP Negeri 1 Tidore Kepulauan with a total of 606 students. In the equating process, the items to use in the item bank were those that had been selected through the item analyses employing both classical theory and modern theory. The parameter estimation was carried out on the basis of the two parameter logistic model from the analysis using the BILOG 3.00 computer program. The data in the form of scores were processed and equated by means of the classical test theory approach (Microcat Iteman) and modern test theory.

The results of the study show that the mathematics test items for the mathematics examinations for Grade VII, VIII, and IX have good quality. Based on classical test theory, the test set for Grade VII consists of 29 items, that for Grade VIII consists of 27 items, and that for Grade IX consists of 30 items. Meanwhile, based on modern test theory, the test set for Grade VII consists of 28 items, that for Grade VIII consists of 27 items, and that for Grade IX consists of 24 items. The equation for equating test sets from Grade VII to Grade VIII is $Q^* = 1.084X - 0.221$, that from Grade VIII to Grade IX is $R^* = 0.922X + 2.965$, and that from Grade VII to Grade IX is $R'^* = 0.998X + 2.775$. The testees in the mathematics examinations for Grades VII, VIII, and IX show that their average ability ranges from 0.265 to 0.58. This means that the testees' ability in the class examinations is still low. The equating of test sets from Grade VII to Grade VIII decreases; the test is easier for Grade VIII students. Likewise, the equating of test sets from Grade VIII to Grade IX decreases; the test is easier for Grade IX students. However, the equating of test sets from Grade VII to Grade IX increases; the test is more difficult for Grade IX students. This is caused by the fact that Grade IX students do not understand the materials for Grade VII. Based on the explanation of the test analyses using two approaches, namely classical test theory and modern test theory (the Marginal Maximum Likelihood model), it can be indicated that in the case of the mathematics subject for the three observed grades there is a difference in the percentages of good and valid items.

Keywords: *vertical equating, learning achievement test, mathematics, item bank*